International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 07 | July - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

TEXT TO SPEECH CONVERSION

[1] SHRUTHI M T, [2] SPOORTHI A S

[1].Assistant professor ,Department of MCA ,BIET ,Davangere[2].Student, Department of MCA, BIET, Davangere

Abstract

Python text to speech translation is the goal of this project. We will develop a Python script to transform text written in human language into voice that sounds like humans. For learning to comprehend a language, I believe listening works better than reading. Many Python APIs are available to convert text to voice. Popular and often referred to as the GTTS API is the Google Text to Speech API. The text to voice API of Google Translate is interfaced with using this Python package.

Among the quickest and simplest approaches to create GUI apps with tkinter is to use the standard GUI Python module. One plays audio files using the playsound module.

We may play a sound file using this module with only one line of code. With so many built-in features, the program is quite simple to use and used to save text files as mp3 files.

This project allows us to input a text message that we want to have spoken, and then we click the play button to hear it said. Initially, we import the modules, then we build the display window and specify functions. Opening the Import Libraries window-function to convert text to speech or voice-function to exit-function to reset define buttons Goggle GTTS API and Python will be used to programmatically transform the text to voice or speech.

1. INTRODUCTION

A fairly user-friendly Python package called GTTS (Google Text-to-Speech) transforms text into audio. One plays audio files using the playsound module.

This module allows just one line of code to play a sound file. To convert text to voice in Python, there are a number of APIs accessible. Among such APIs is the gTTS API, or Google Text to Speech API.

The extremely user-friendly gTTS program turns input text into audio that may be saved as an mp3 file. Users of this project may quickly turn emails, articles, or any other textual material into voice for accessibility or on-the-go listening. Though this technology has several drawbacks, the text-to-speech project is a major development in the areas of accessibility and digital communication.

2. RELATED WORK

Shashank Tripathi suggests in N. K. P. S. a method that allows blind and visually impaired persons to utilize email services as effectively as a regular user. The system runs on STT and TTS processes and is virtually mouse and keyboard independent. The user identification is also authenticated via face recognition. Some techniques for speech representation and categorization were proposed. They also used many feature extraction methods along with database performance and assessment. Analyzing the many issues with Automatic Speech Recognition, the authors suggested solutions. They examine the AI Approach, the Pattern Recognition Approach, and the Acoustic Phonetic Approach among other approaches to voice recognition. Shashank Tripathi suggests in N. K. P. S. that the person utilizing the system should train it.

Both a high command count and word recognition accuracy of over 95% are achievable by these systems.

The disadvantage of this method is that the system only reacts precisely to the person who taught the system. Software for home computers most often uses this method.

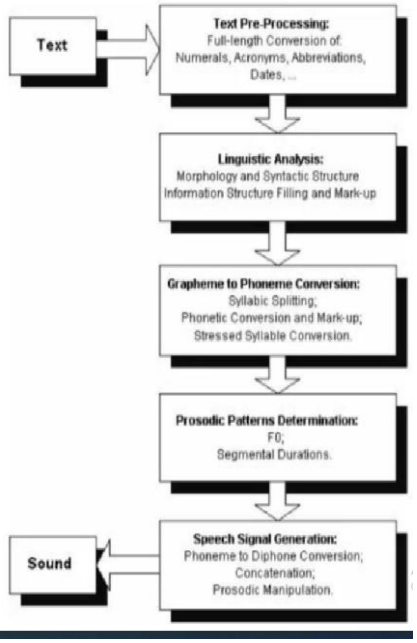
Hirschberg, Sproat, Olive, J.P., and Van Santen proposed that the capacity of a speech synthesizer to be comprehended and its resemblance to the human voice determine its quality. Written works may be listened to on a home computer by anyone with reading or visual impairments thanks to an understandable text-to-speech application. A front-end plus a back-end comprise a text-to-speech system, often known as a "engine".

Ramesh Kagalkar proposed in Kaveri Kamble the development of a TTS system for regional languages like Hindi. Two primary processes of the system are speech generation and text pre-processing. We investigate a concatenative synthesis- based method to extract the speech from the text. Furthermore included is a spellchecker module to verify that words in local tongues like Hindi are spelled correctly. After evaluating the many publications, we have come to the

conclusion that the fields of text-to-speech and speech-to-text conversion still have a long way to go. The following part contains an analysis of many TTS and TTS synthesis techniques.

According to S. R. Mache, the test-to-speech synthesizer has been evolving quickly throughout the last several years to take on its present form. TTS is best served by the approaches of concatenative synthesis, articulator, and formant. Some research groups are even developing Text-to-Speech in regional languages such Marathi, Hindi, Telugu, Punjabi, Kannada, and so on in India. TSS synthesis may be greatly improved to get a respectable level of natural and emotional aspect.

3. METHODOLOGY



a.

a. Problem Synopsis

This part gives a succinct and understandable summary of the issue; it should also include the facts that were utilized in the project as well as the research questions and theories that served as project guidance. This issue statement need to make it very evident what the research paper aims to solve and how the project will tackle it.

b. Practice:

Platform agnostic pyttsx3 library is used to develop the Text-to-Speech converter. Text-to-speech conversion using this library has the main benefit of working offline. Still, Pyttsx is limited to Python 2.x. Thus, pyttsx3 will be shown, which has been updated to function with Python 3.x as well as Python 2.x.

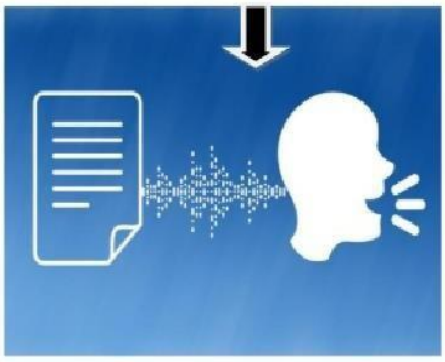
Using voice synthesis methods, the artificial speech will have a naturalness character.

2 The fundamental unit for voice synthesis might be the English language process.

3 Phoneme-based speech database development for the English language will be done.

Four phonemes will be searched in the database and the sounds of the matching phonemes will be concatenated to produce synthesised output.

4. RESULTS



5. FUTURE ENHANCEMENT

There will be many and interesting improvements to the text to voice converter project in the future. Adding other languages to the system might be one approach to improve it. The project's present support for a small number of languages may make it less helpful for users who speak other languages or in specific areas. A greater variety of users may find the project more inclusive and accessible if it were to be expanded to include other languages. The addition of more sophisticated artificial intelligence and natural language processing technologies might be another improvement. More accuracy and responsiveness would result from the technology being able to comprehend and interpret the subtleties of human speech.

To further facilitate more natural and intuitive user interaction with the system, it may also be linked with other technologies like virtual assistants. At last, the project may be extended to incorporate new functions and capabilities, including the capacity to alter the voice and accent utilized for the speech output or to preserve and save earlier speech conversions. These improvements would increase the system's flexibility and suitability for a larger number of uses, and they may even spur the creation of novel and cutting edge speech synthesis technologies.

6. CONCLUSION

In summary, the text-to-speech converter project has been successfully finished and has given rise to a practical and easy-to-use tool for doing this. The initiative sought to provide those who find reading difficult or who would rather listen to material than read it a solution. Through the use of the Pyttsx3 library to convert the text to voice and the provision of an easy-

to-use user interface, the project has succeeded. All things considered, the initiative has been successful and may be improved with more functionality like voice customisation and language translation.

7. REFERENCES

1. Review on Text-To-Speech Synthesizer,

S. R. Mache, International Journal of Advanced Research in Computer and Communication Engineering, 2015

2. N. K. P. S. Shashank Tripathi, "Voice based Email System for Differently abledand Visually Impaired," International Journal of Engineering Research & Technology, 2019

3. Rubin, P.; Baer, T.; Mermelstein, P. (1981). A perceptual research articulatorysynthesizer. American Journal of Acoustics, 70, 321–328.

4. Van Santen, J.P.H., Olive, J.P., Sproat,and Hirschberg, J. (1997). Speech Synthesis Advancement. Heidelberger.

5. Fifth In T.A. Sebeok (Ed.), Current Trends in Linguistics, Vol. 12, (1974), p.2451–2487, Mattingly I. G., Speech Synthesis for Phonetic and PhonologicalModels.

6. A Review: Text to Speech Conversationfor Hindi Language, Kaveri Kamble and Ramesh Kagalkar, International Journal ofScience and Research (IJSR), 2